CHAPTER 5

AN IMPROVED REAL TIME IMAGE DETECTION SYSTEM FOR ELEPHANT INTRUSION ALONG THE FOREST BORDER AREAS

5.1 INTRODUCTION

Human-elephant conflict is a major problem leading to crop damage, human death and injuries caused by elephants, and elephants being killed by human. In this work an automated unsupervised Elephant Image Detection System (EIDS) as a solution to human elephant conflict in the context of elephant conservation is proposed. The elephant image captured in the forest border areas are sent to a base station via a radio frequency (RF) network. The received image is decomposed using Haar wavelet to obtain multilevel wavelet coefficients, with which image feature extraction and similarity match between the elephant query image and the data base image is performed using image vision algorithms. A GSM message is sent to the forest officials indicating that an elephant has been detected in the forest border and is approaching human habitat. An optimized distance metric to improve the image retrieval time from the database is proposed. The optimized distance metric is compared with the popular Euclidean and Manhattan distance methods. The proposed optimized distance metric retrieves more images with lesser retrieval time than the other distance metrics which makes the proposed method more efficient and reliable.

5.2 IMAGE CAPTURE HARDWARE MODEL

Elephants walking in the sensing range of the geophones produce vibration which is converted to an electrical signal processed in an embedded controller and a SMS is sent to the forest officials for necessary action. In such intrusion detection systems there is a possibility of insufficient vibrations sensed by the geophones as a consequence of weather conditions like rain, soil moisture. Due to this it may miss the event i.e. elephant movement. An image processing based approach is developed as a solution to the above said problem to identify an intruding elephant in to human living areas. The elephant comes out of the forest in certain pockets to enter in to human living areas for getting food and water. The cameras mounted on towers or trees capture the image of the intruding elephant and send the image to the base station using RF network. The received image is processed at the base station and compared with the stored data base images of elephants. The hardware module has two sections the transmitter and receiver as shown in Figure 5.1.

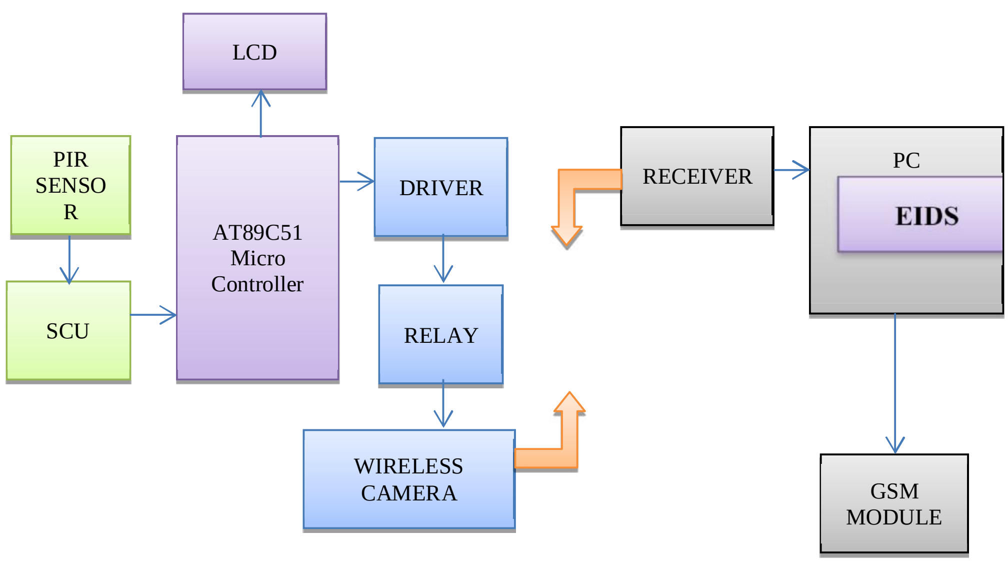

Figure 5.1 Image Capture Hardware Block Diagram

5.2.1 Transmitter Section

The transmitter section setup consists of wireless camera, PIR sensor and Atmega embedded controller as shown in Figure 5.2. The PIR (Passive Infra-Red) Sensor is a pyro electric sensor device that detects motion by measuring changes in the infrared (heat) levels emitted by surrounding objects. When elephant motion is detected the PIR sensor provides a high signal on its output pin. This logic signal can be read by the embedded controller that switches the camera to the capture mode; the camera captures the image over a 20 meter distance. The captured video is transmitted to the base station via RF network. The elephant pockets in the corridors are identified and these setups are installed to monitor the movement of the pachyderms.

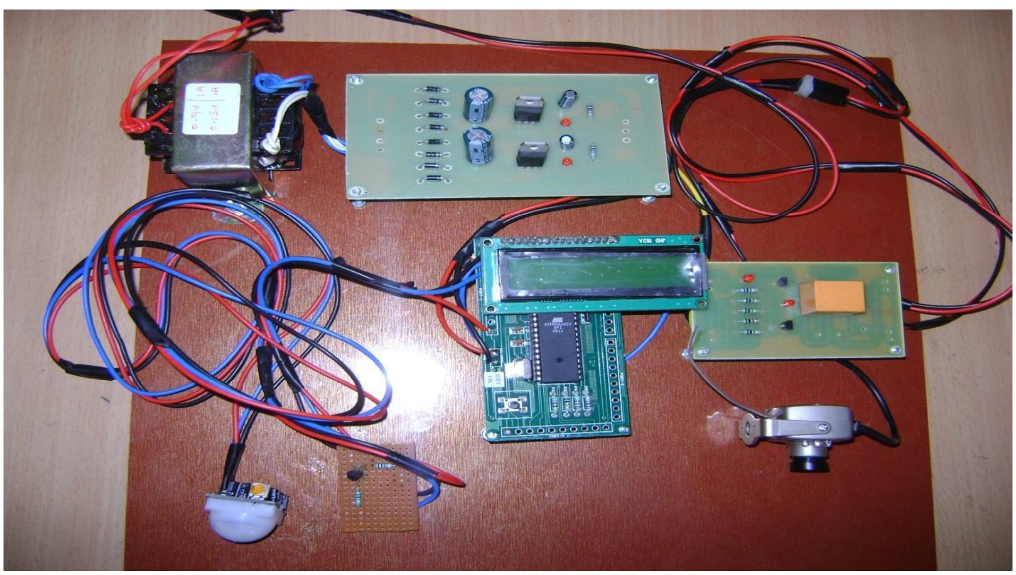

Figure 5.2 Image Capture Transmitter module

5.2.2 Receiver Section

The RF transmitter transmits the captured video to the RF receiver. A TV tuner card is used as a receiver which is an adjustable wireless video receiver as shown in Figure 5.3. The receiver range is around 150 - 200 feet. System: PAL/CCIR NESC/EIA and the receiving frequency: 900MHz-1200MHz. The video is converted to snap shots every five seconds and stored in the memory. The snapshot from the video is taken every 5 Seconds as shown in Figure 5.4.

Figure 5.3 Wireless camera and Receiver

The stored image is updated in the data base every 5 seconds and is identified as the query image. The query image is compared with the database image. This updated image is added to the database. When image match occurs a SMS is sent to the forest officials through the GSM transceiver connected with the PC.

Figure 5.4 Image Capture Receiver module

Wireless security cameras function best when there is a clear line of sight between the camera(s) and the receiver. With a clear line of sight, digital wireless cameras typically have a range between 250 to 450 feet outdoor. The range is limited to 100 to 150 feet in indoor environment . The signal range varies depending on the type of building materials and/or objects the wireless signal must pass through.

5.2.3 GSM Module

The GSM Modem can accept any GSM network operator SIM card and act just like a mobile phone with its own unique phone number as shown in Figure 5.5. Advantage of using this modem is its RS232 port can be used to communicate and develop embedded applications. Applications like SMS Control, data transfer, remote control and logging can be developed easily. The modem can either be connected to PC serial port directly or to any microcontroller. It can be used to send and receive SMS or make/receive voice calls. It can also be used in GPRS mode to connect to internet and do many applications for data logging and control. In GPRS mode you can also connect to any remote FTP server and upload files for data logging. This GSM modem is a highly flexible plug and play quad band GSM modem for direct and easy integration to RS232 applications. It supports features like Voice, SMS, Data/Fax, GPRS and has an integrated TCP/IP stack.

Figure 5.5 GSM module

5.3 ELEPHANT IMAGE DETECTION SYSTEM (EIDS)

Elephant Image Detection System (EIDS) algorithm is developed in this work. A database of 114 images is created by capturing 2 elephants in different postures of elephants in the Sadivail elephant camp in Coimbatore, South India. The elephant database images are feature extracted using Harr wavelet technique and clustered into groups by K-means clustering. A similarity comparison is made by determining the number of significant coefficients in common between the query signature and the signatures of database using F Norm theory. The searched elephant images are then arranged according to the similarity value obtained in a decreasing order. If the matched images are more than 5, an “elephant detected” message is sent through the GSM to the mobile phone numbers stored in the system. Once the procedure is complete, the system captures the next image and precedes the same steps to detect elephants. The following flow graph in Figure 5.6 shows the proposed algorithm for EIDS

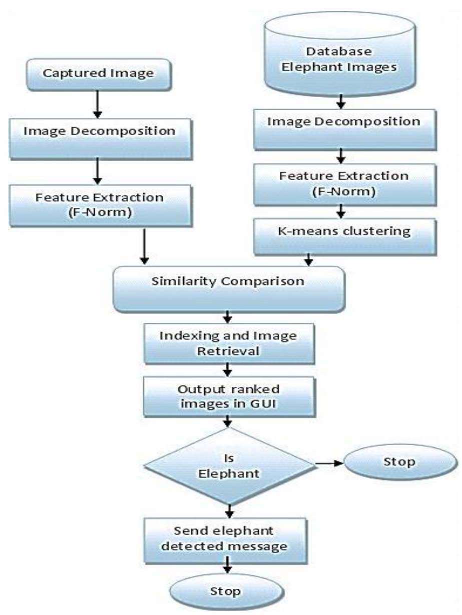

Figure 5.6 EIDS Flow Chart

5.4 HARR WAVELET DECOMPOSITION

The Harr wavelet decomposition of elephant image in RGB color space is represented at multiple scales (Porwik 2004). The Haar wavelet decomposition is computed by iterating difference d_i and averaging a_i between odd and even samples s_i of the elephant image. Averaging and differencing the elephant image elements are done as in equation (5.1),

$$a_i = \frac{S_i + S_{i+1}}{2} \qquad di = \frac{S_i - S_{i+1}}{2} \tag{5.1}$$

If an elephant image data set $S_1, S_2 \ldots S_{N-1}$ contains N elements, there will be N/2 averages and N/2 wavelet coefficient values (Arivazhagan 2003). The averages are stored in the upper half of the N element array and the difference coefficients are stored in the lower half of the array. The averages become the input for the next step in the wavelet computation, for iteration i+1, $N_i = N_{i/2}$. The recursive iterations continue until a single average and a single difference are calculated (Latha 2007). The scaling and wavelet values are represented by h_i and g_i respectively as given in Equation a (5.2) and (5.3).

$$h_0 = 0.5, \qquad h_1 = 0.5 \tag{5.2}$$

$$g_0 = 0.5, \qquad g_1 = -0.5 \tag{5.3}$$

The Haar transform are shown in equation (5.4) in matrix form.

$$\begin{matrix} h_0 & h_1 & 0 & 0 & \ldots \\ g_0 & g_1 & 0 & 0 & \ldots \\ 0 & 0 & h_0 & h_1 & \ldots \\ 0 & 0 & g_0 & g_1 & \ldots \end{matrix} \tag{5.4}$$

The Haar transform for an eight element signal is shown in equation (5.5). Here signal is multiplied by the forward transform matrix A.

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ d_0 \\ d_1 \\ d_2 \\ d_3 \end{pmatrix} = A \bullet \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \end{pmatrix} \tag{5.5}$$

where,

$$A = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & -0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & -0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & -0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.5 \end{bmatrix}$$

Since the columns of the A are orthogonal to each other each of these matrices is invertible with respect to A. The elephant database images are decomposed into multi-level coefficients. After decomposition, feature vectors for all the elephant images in the database are obtained using F-norm theory (Manimegalai 2010) as given in equations (5.6) and (5.7). Every image is considered as a square matrix. A is a square matrix and A_i is its i_{th} order sub matrix. The F-norm of A_i is given as:

$$||A_i||_F = \left[\sum_{K=1}^{i} \sum_{I=1}^{i} |a_{KI}|^2\right]^{1/2} \tag{5.6}$$

Let**,**

$$\Delta A_i = ||A_i||_F - ||A_{i-1}||_F \text{ and } ||A_0||_F = 0$$

The feature vector of A is defined as:

$$V_{AF=} \left\{\Delta A_{1,}\Delta A_{2}, \dots \dots \dots \Delta A_{n}\right\} \quad (5.7)$$

Vector elements in the feature vector are represented by ΔAi and ΔBi. The similarity between the two images is given by the following similarity criteria ((Latha 2007). Let α_i be the similarity of ΔAi and ΔBi.

$$\alpha_i = \begin{cases} \min(\Delta A_i, \Delta B_i)/\max(\Delta A_i, \Delta B_i), & \Delta A_i \neq 0 \text{ or } \Delta B_i \neq 0 \\ 1, & \Delta A_i = \Delta B_i = 0 \end{cases} \quad (5.8)$$

Similarity between the two images lies in between 0 and 1i.e. ($0 \leq \alpha \leq 1$). The images in the data base are arranged according to the similarity match with the query image.

5.4.1 K-Means Clustering Algorithm

Clustering is a process of grouping the similar objects from a given data set. The most popular and reliable clustering algorithm is the K means clustering algorithm that classifies the input data points into multiple classes based on their inherent distance from each other. Let S = {S_i, i= 1, 2...N} be the n-dimensional data points to be clustered into a set of K-Clusters, C = {$C_{1,}$ C_2, $C_{3\dots}C_K$} (Doreswamy 2012). From the given elephant data set X $=\{x_1 \dots\dots, x_N\}$, x_n □ E^d. The M-clustering problem aims at partitioning the elephant data set into M disjoint subsets (clusters) $C_1 \dots C_M$. Based on this criterion the elephant images are grouped depending on the cluster centers $m_1 \dots m_M$ as given below

$$E\ (m_1 \dots\dots m_M) = \sum_{i=1}^{M} \sum_{k=1}^{M} I(x_i \in C_k)\ \|x_i - m_k\|^2 \quad (5.9)$$

where, I (X) = 1 if X is true and 0 otherwise.

5.5 PROPOSED OPTIMIZED DISTANCE METRIC

In this work a novel distance metric called optimized distance measure integrated with K-means clustering algorithm to improve retrieval time is proposed. The distance metrics are used in the work for (i) finding similarity between two images and (ii) ordering a set of images based on their distances from a given image. In many image retrieval systems Euclidean (Doreswamy 2012) and Manhattan (Vadivel et al 2003) are the popular distance measure algorithms used. A study on the above two similarity measures is carried out and a new distance method called optimized distance measure is proposed. It is observed that the proposed method retrieved more images with faster retrieval rate than the other two methods.

5.5.1 Euclidean Distance Method

The Euclidean distance method (Doreswamy 2012) is suitable for the correlation between quantitative, continuous variables and not suitable for ordinal data and is given as

$$D = \sqrt{\left(R_c - R_g\right)^2 + \left(G_c - G_g\right)^2 + \left(B_c - B_g\right)^2} \qquad (5.10)$$

where (R_c, G_c, B_c) are centroids and (R_g, G_g, B_g) are the pixel or data points. Most of the time is spent to calculate the square root so it is basically time consuming. The Euclidian scheme produced 20 retrieved images with retrieval time of 21.544 seconds and the retrieval rate per image is 1.077 seconds as shown in Figure 5.7.

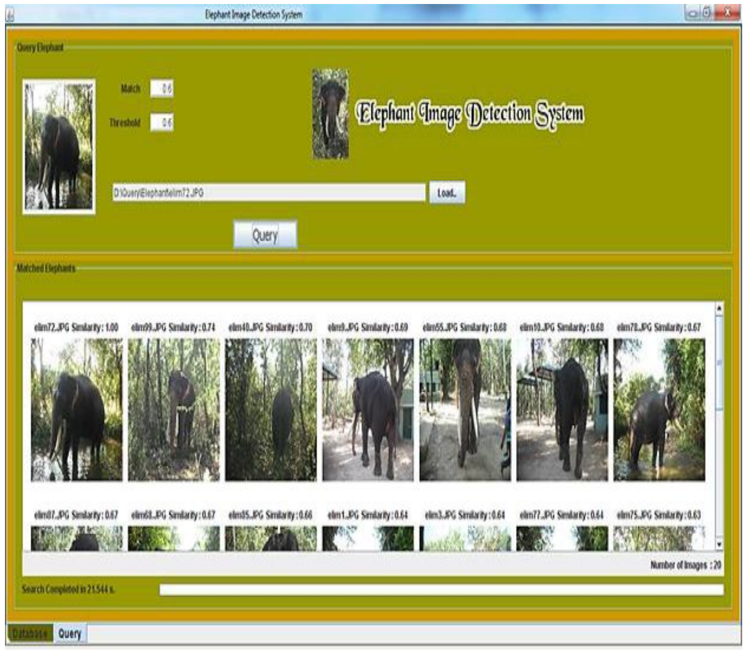

Figure 5.7 Elephant image detection system using Euclidian distance method

5.5.2 Manhattan Distance Method

The Manhattan distance is the absolute sum of the horizontal and vertical components of the image data matrix. This is essentially a consequence of being forced to adhere to single-axis movement, (i.e.) one cannot move diagonally in more than one axis simultaneously and the distance is given by

$$\text{Distance d} = \sum_{s_i \in c_k} \left\| s_i - \mu_k \right\| \qquad (5.11)$$

Whenever each pair is in nonempty intersection, there exists an intersection point for the whole collection; therefore, the Manhattan distance forms an injective metric space. The Manhattan scheme produced 22 retrieved images with retrieval time of 15.067 seconds and the retrieval rate per image is 0.684 seconds as shown in Figure 5.8.

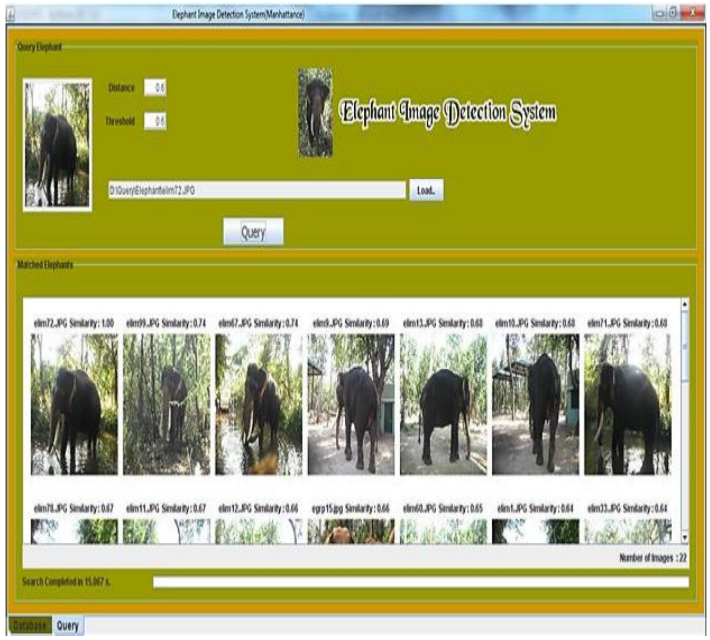

Figure 5.8 Elephant image detection system using Manhattan distance method

5.5.3 Optimized Distance Method

In this proposed optimized distance method for the given query image only the distance related cluster is searched. The optimized distance metric is given as

$$D = \left\| \sum_{i=0}^{n} \left[\left(R_c - R_g\right)^3 + \left(G_c - G_g\right)^3 + \left(B_c - B_g\right)^3 \right] \right\| \tag{5.12}$$

The optimized distance method computes the cube power of the distance between the centroid and color pixel points of the three colors and determines the summation of all the added values. The modulus of the whole summed values is calculated to get the distance value. For the proposed optimized distance metric the images retrieved were 27 with retrieval time of 15.028 seconds and the retrieval rate per image is 0.556 seconds and shown in Figure 5.9.

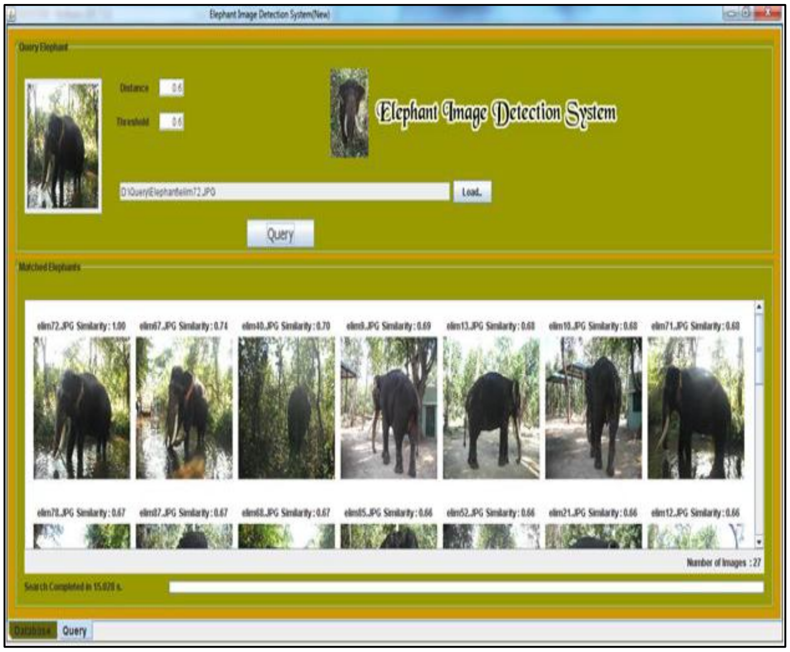

Figure 5.9 Retrieved Images for Optimized Distance Metric

5.6 RESULTS AND DISCUSSION

5.6.1 Field Observations

Field observations are carried out in the forest border areas in Sadivail elephant camp as shown in Figure 5.10. The hardware setup was arranged to capture the image of elephants. The wireless camera was mounted on a wooden stick and the camera was battery powered with 12V. The blue circle indicates the camera mounted on a stick. Using RF receiver the video images are received and is converted to image frames using camcorder software. The elephant image frames are stored in memory and updated every 5 seconds. The proposed method is implemented in Matlab and Java programming language.

Figure 5.10 Field Observations in Sadivail Elephant Camp

Figure 5.11 shows the on line Elephant Image Detection System window in which 12 elephant images are retrieved in 6.33 seconds. The elephant images are arranged in the order of similarity value obtained. As the retrieved images are more than 5, a GSM message "Elephant Detected" is sent to the forest officials using the AT command.

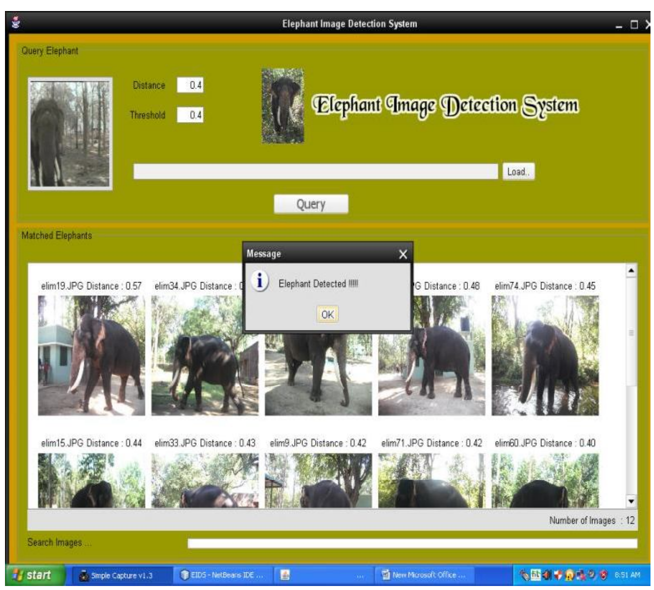

Figure 5.11 Retrieved Images - On line Elephant Image Detection

5.6.2 Wavelet Analysis

In this work, Haar wavelet decomposition on the raw elephant image by determining the scaling coefficient and finding the largest wavelet coefficient is performed. The scaling coefficient is stored along with the difference and location (*i*, *j*) of each wavelet coefficient for every image (Arivazhagan 2003). The 3 level decomposed query elephant image is shown in Figure 5.12. Using the 3 level wavelet decomposition, the highest and informative elephant image features were extracted from the coefficients. These features are used during the process of the query and database image comparison of the elephants. The Harr wavelet transform of the elephant is calculated by passing it through a series of filters (high and low pass filters) and then down-sampled, as shown in Figure 5.13.

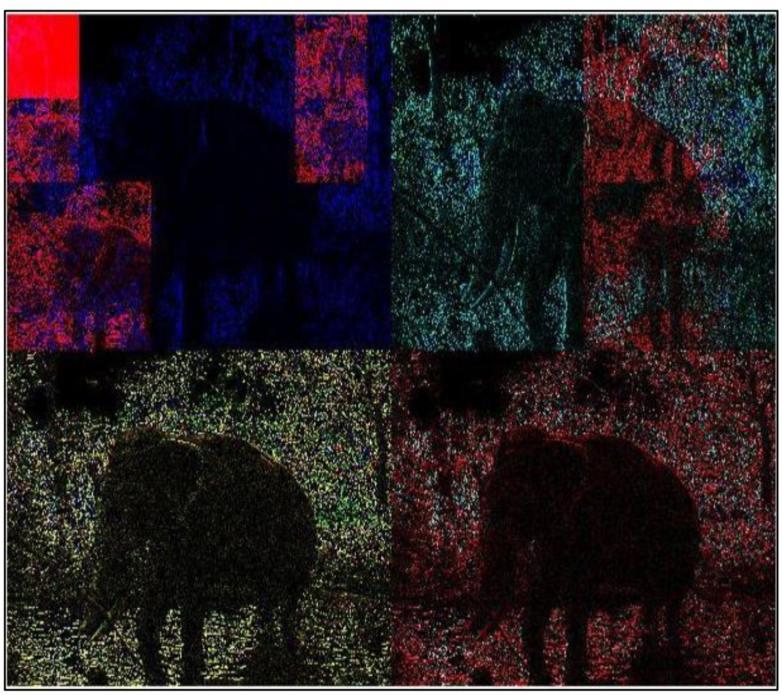

Figure 5.12 Level Decomposed Elephant Image

Figure 5.13 Down-Sampled Decomposed Elephant Image

At each level, the elephant image is decomposed into low and high frequencies, and this decomposition halves the resolution since only half the number of samples are retained to characterize the entire image. The Harr wavelet leads to a decomposition of approximation coefficients at level j in to four components and at level j+1. Due to successive down sampling by 2, the image length must be a power of 2, or a multiple of a power of 2, and the length of the image determines the maximum levels in to which the elephant image coefficients can be decomposed. The Harr wavelet coefficients of different species are plotted in graph as shown in Figure 5.14. The Harr wavelet coefficient of each species varies with the elephant image and can be distinguished from other animals.

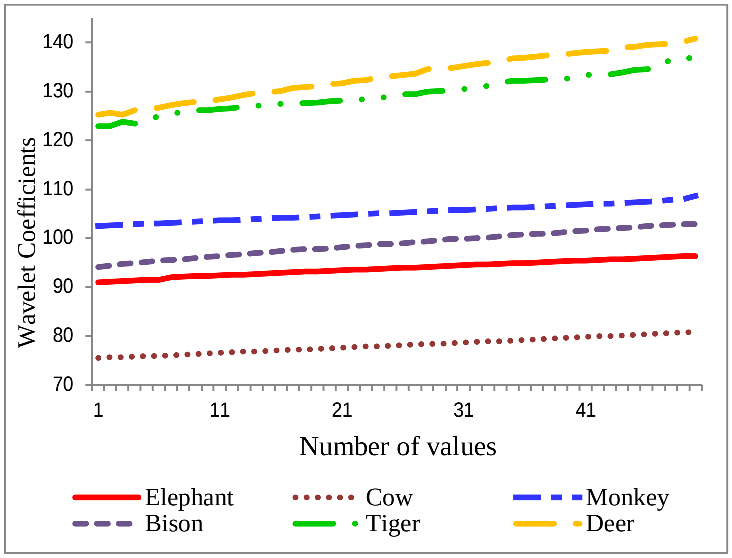

Figure.5.14 Wavelet Coefficients for Different Species

The value of wavelet coefficients of elephant and bison are closer because the major color is black for both the species and the value lies in between 90 to 95. The coefficients for tiger and deer are in the higher band in between 120 to 130. The obtained elephant coefficient is averaged to get the threshold value. We fix 0.6 as threshold value to obtain the similar elephant images from the database image for the given query image. The EIDS system is tested offline with elephant and non-elephant images. All the images used in this work are in the dimensions of 3648 x 2763. We also tested the system for group of elephant images and it is shown in the Figure 5.15.

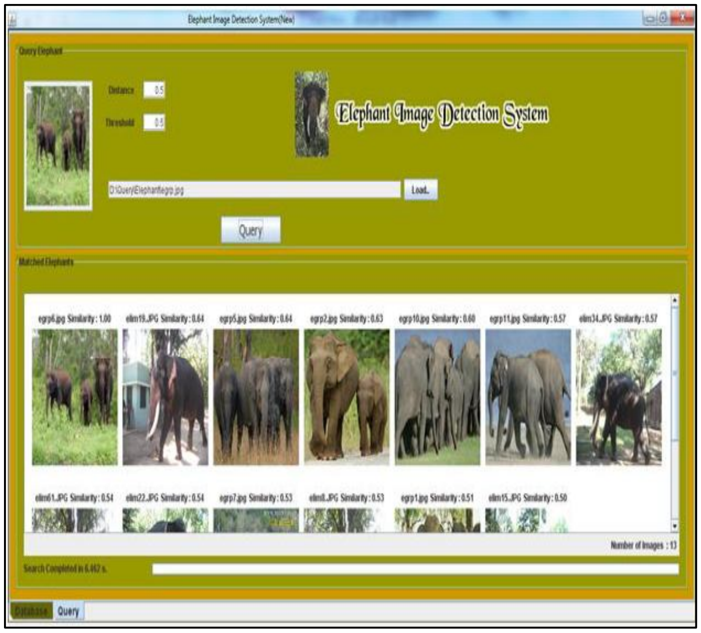

Figure.5.15 Retrieved Elephant Group Image

The elephant normally moves in herds in the forest borders during the period of migration. For the given query image the system retrieved 13 images in 6.462 seconds. We tested the system with group elephant images of different sizes and postures. Non-elephant images like that of bison, bear, deer, monkey, and human are most commonly seen around forest border areas. The non-elephant images are given as query, which produced zero image search result. We tested with a bison image as it is of similar color texture of elephants and the system retrieved zero search images as shown in Figure 5.16. On zero search result, no alert is made, hence GSM message is not sent.

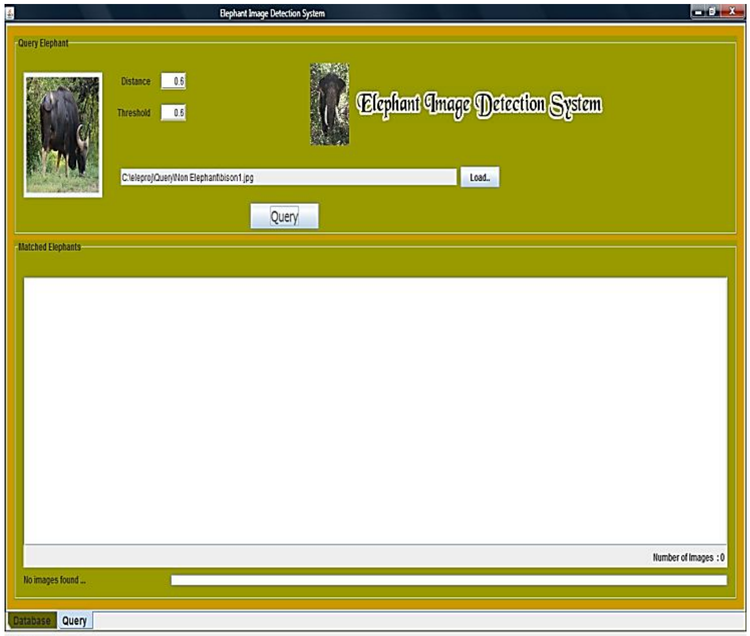

Figure 5.16 Retrieval for Non elephant query image

5.7 COMPARISON BETWEEN THE DISTANCE METHODS

A comparison between the distance methods is carried out in this work and the results are shown in Figure 5.17. It is observed that the proposed method retrieves more images in lesser time compared to the other two methods. The optimized retrieval rate improvement over other distance metric is 18%. To assess the retrieval effectiveness, 10 query elephant images are selected and tested. The average retrieval time per query elephant image is calculated for all the three distance measures and the results are shown in Fig 5.18. The retrieval time per image is less for the proposed distance metric compared to other methods. Due to reduction of computational time and higher image retrieval rate, the time to react on elephant intrusion is improved.

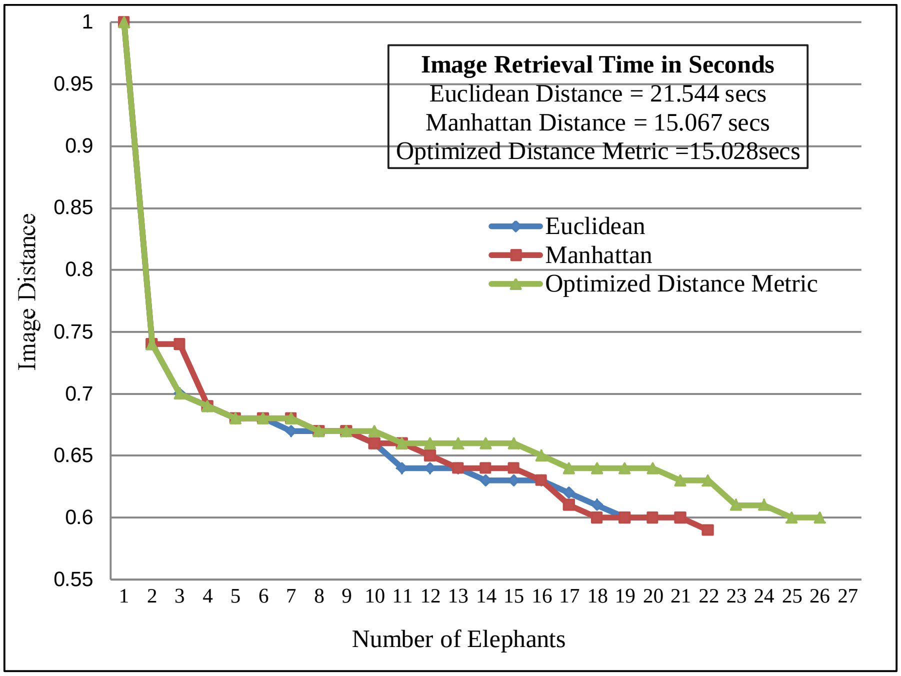

Figure 5.17 Image Distance Metric Comparison

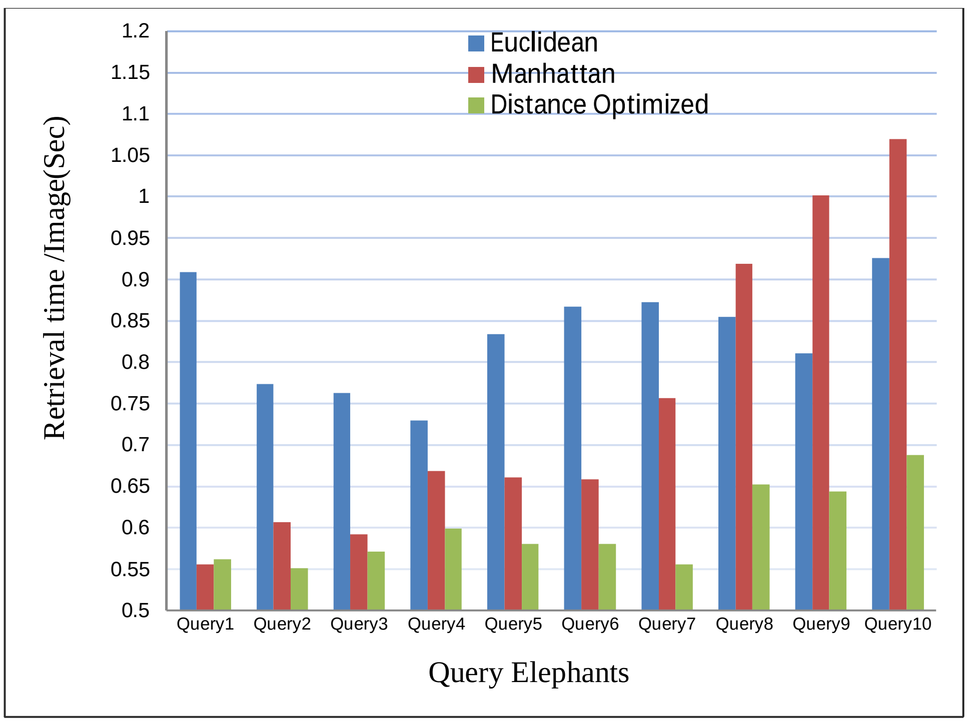

Figure 5.18 Image Retrieval Rate Comparison

The Recall rate is defined as the ratio of the number of relevant (same shape and position) retrieved images to the total number of images in the data base.

$$\text{Recall rate} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images in database}}$$

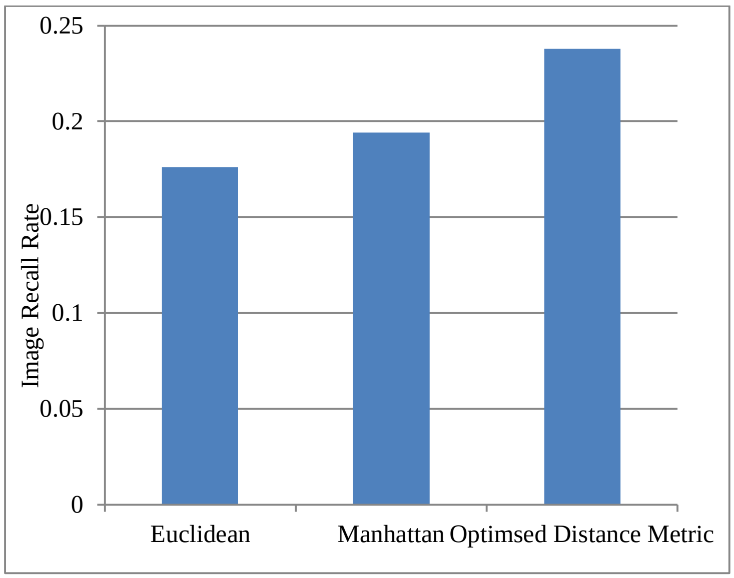

Figure 5.19 Elephant Image Recall Rate Comparison

A comparison is made on image recall rate for the three methods as shown in Figure 5.19 in which the optimized distance metric is 16% better than Manhattan and 18.5 % better than Euclidean method. The algorithm is tested by varying the cluster formation. Table 5.1 shows the variation of number of images retrieved, retrieval time and retrieval rate per cluster.

Table 5.1. Clustering Database Images for Different K Values

K	No of retrieved images	Retrieval time for images	Retrieval rate per image
2	14	10.884	0.777
3	12	7.604	0.6336
4	12	7.242	0.6032
5	9	4.949	0.5498

When 2 clusters is fixed, 14 images are retrieved and the retrieval time for the 14 images is 10.884 and the retrieval rate per image is 0.777.

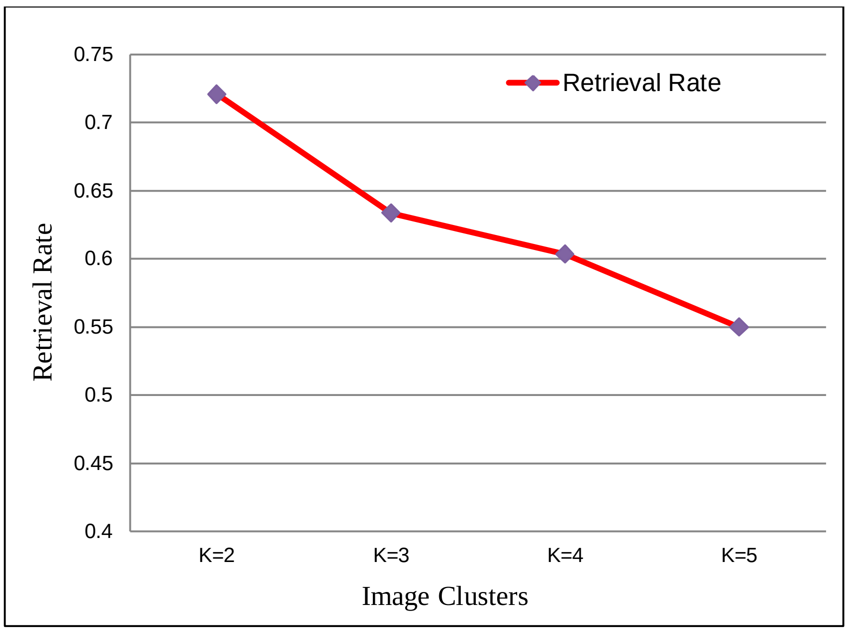

Figure.5.20 Image Retrieval Rate per Clusters

The cluster value K is varied from 2 to 5 and recorded the corresponding number of retrieved images and the retrieval rate per image using the database for the query image given. It is seen that for small number of clusters more images are retrieved and for higher number of clusters the retrieved images were less as shown in Figure 5.20. So for the analysis we fix 2 clusters in order to retrieve more images from the database.

A comparison was made between off-line and on-line elephant image detection system. In off line system the query images are given manually and the performance of the algorithm developed is tested. In the online system the query images are obtained from the camera lively and the captured image is given as a query image then the software runs for retrieving elephant images from the database. Once 5 images are retrieved then an alert is sent. In the online system for the threshold values 0.4 to 0.6 the number of images retrieved is 5 to 12, for off-line system it produced 17 to 22 images. This variation is due to the camera posture and image capture in frame in the online system. The required number of 5 images is achieved in on-line system with the proposed optimized distance method makes the system efficient and reliable. The result comparison between the online and offline is recorded in the Table 5.2.

Table 5.2. Comparison between the On-Line and Off-Line Observations.

Image Threshold	Image Distance		Retrieved Images	
	Offline	Online	Off line	On line
0.4	1.0 – 0.49	0.57 -0.36	22	12
0.5	1.0 – 0.49	0.44 -0.34	22	6
0.6	1.0 - 0.59	0.47 -0.38	17	5
0.7	1.0 – 0.74	0	2	0

The following figures are the online images obtained by varying the image threshold in the EIDS system.

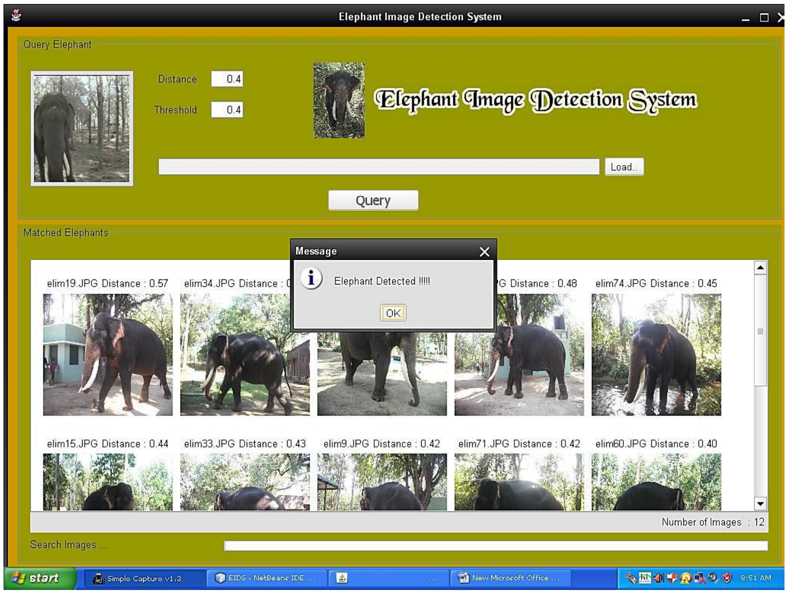

Figure 5.21 Online Image Retrieval for threshold 0.4.

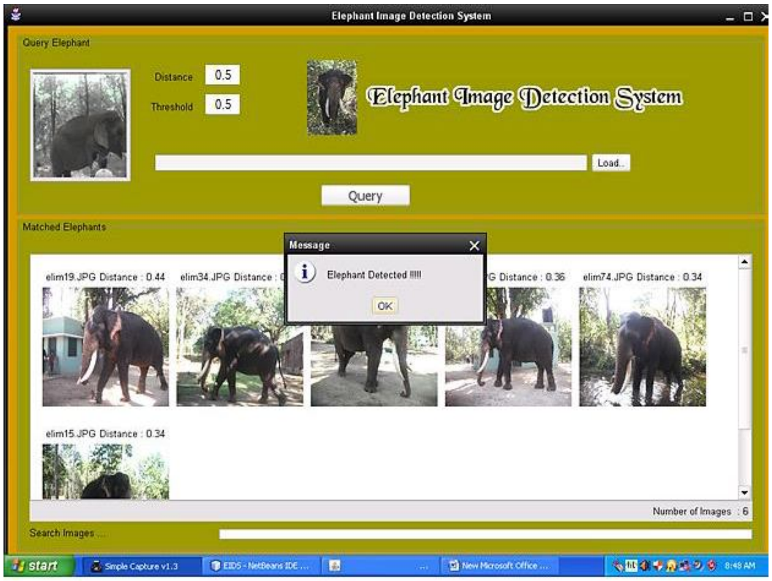

Figure 5.22 Online Image Retrieval for threshold 0.5.

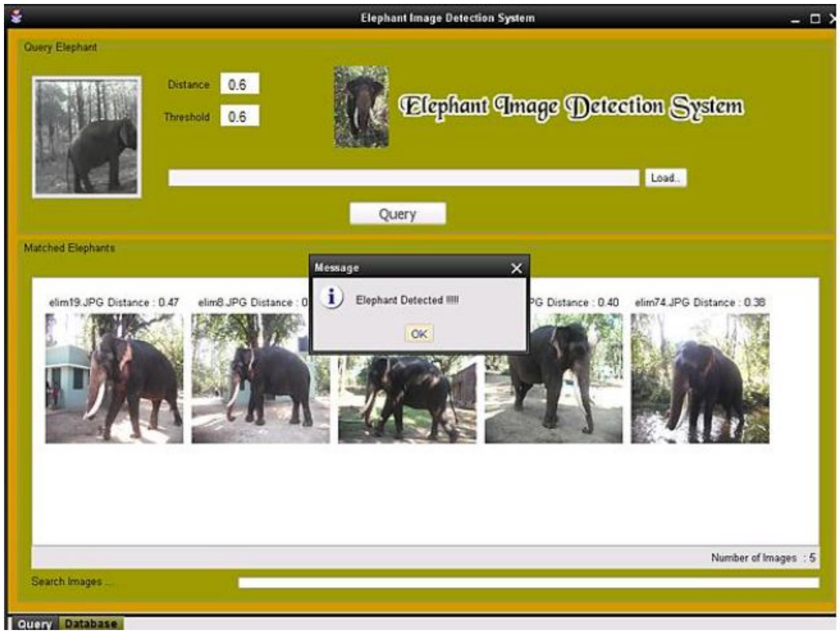

Figure 5.23 Online Image Retrieval for threshold 0.6.

Figure 5.24 Online Image Retrieval Rate for threshold 0.7.

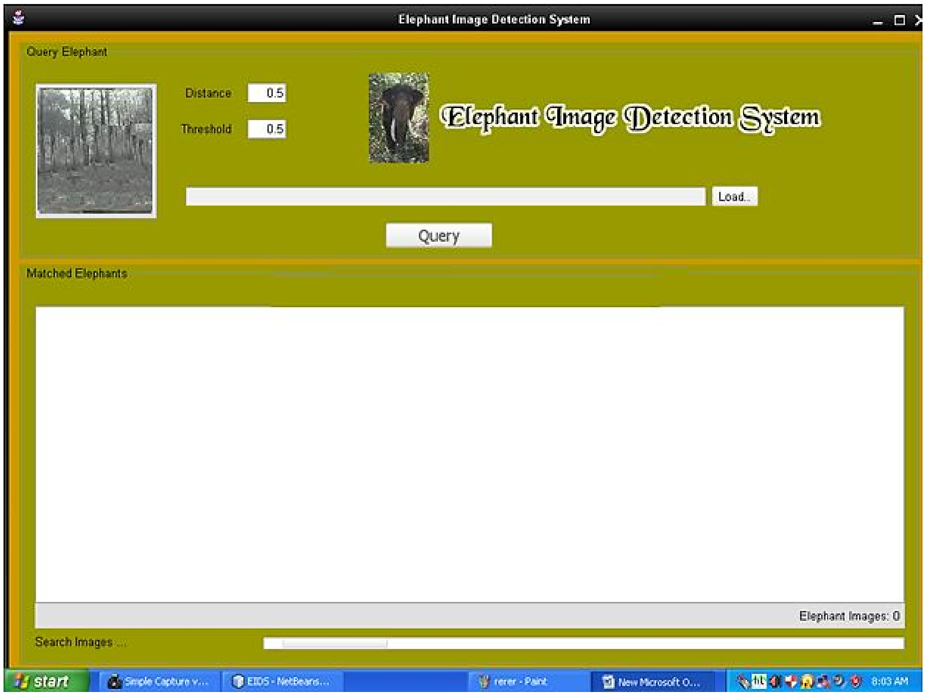

Figure 5.25 Image retrieval when no elephant is present

5.9 CONCLUSION

In conclusion, the findings of our work contribute to elephant conservation issues. The work provides solutions to human elephant conflict. The study provides insights to protect elephants from human activities and reduces the work effort of forest officials. The real time elephant identification system provides solutions to the problem of human elephant conflict and provides solution for unsupervised process of individual species identification specifically for elephants. The system is completely automated; the strength of this approach stems from the ability to narrow the collection of potential matches in the database with the query image. Optimal results for automated identification of individual elephants are obtained with the algorithm developed and is used to rank the most likely matches, followed by final supervised visual identification with an early warning sent to the forest officials about the arrival of elephants from the forest borders in to the human habitat. The real time automated approach minimizes the manual work which is not possible all the time because it is difficult to monitor the presence of elephants manually when the herds take a march towards the forest borders. More importantly, our results demonstrate the importance of certainty in identifying approaching elephants in to human living areas and providing early warning about the elephant entry in to the human habitat. We therefore recommend the use of the real time image processing technique to identify approaching individual elephant and group of elephants. The system can be laid in forest border areas. We also solve the traits for distinguishing species where animals are differentiated on the basis of Harr wavelet color attributes. In the context of elephant conservation the real time automated image processing system can be used for the elephant ecology learning, population monitoring and elephant habitat usage. The system can also be deployed along forest border migration routes or at water holes and food plantation areas for elephant tracking and monitoring.

In this work an Elephant Image Detection System via wavelet decomposition of images, followed by feature extraction and similarity match under F-norm theory is proposed. The retrieval performance of optimized distance metric based K-means clustering is compared with the existing techniques like Euclidian distance and Manhattan distance. It turns out that optimized distance metric calculation has high retrieval rate as well as enough recall rate. Field observations show that the proposed method can be used as an effective scheme to detect elephants in the forest border areas even in the presence of different species. This system has been rigorously tested throughout the various phases of the project and found to be efficient compared to the existing systems.